Teaching Machine Learning using Data for Good

BIRS 24w5284: Statistical Aspects of Trustworthy Machine Learning
February 13 2024

Dr. Lloyd T. Elliott
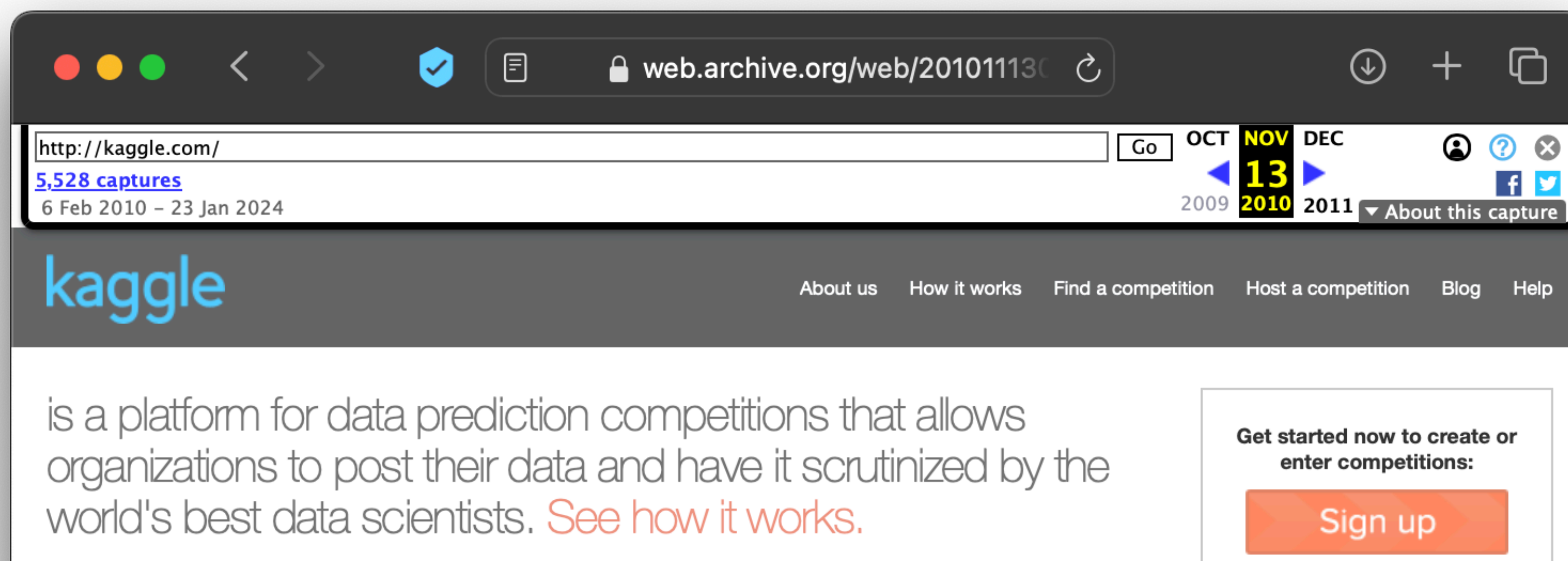Simon Fraser University

Hackathons!

- We compete to find the best solution to a case study (can be quantitative)

  - Benefit to student

  - Benefit to organizer

  - Benefit to commons/public good

Hackathons!

- Popularized by the Netflix Challenge

  - $1m prize to improve 'Cinematch'

  - ~100m ratings, ~1/2m users, ~20k movies

  - Test set of ~2m ratings are held out

  - Teams upload predictions of ratings

 - Origin of collaborative filtering (Mnih et al. 2008)

Rank	Team Name	Best Test Score	% Improvement	Best Submit Time
Grand Prize - RMSE = 0.8567 - Winning Team: BellKor's Pragmatic Chaos				
1	BellKor's Pragmatic Chaos	0.8567	10.06	2009-07-26 18:18:28
2	The Ensemble	0.8567	10.06	2009-07-26 18:38:22
3	Grand Prize Team	0.8582	9.90	2009-07-10 21:24:40
4	Opera Solutions and Vandelay United	0.8588	9.84	2009-07-10 01:12:31
5	Vandelay Industries !	0.8591	9.81	2009-07-10 00:32:20
⋮				
111	ML@UToronto A	0.8793	7.69	2007-09-14 18:56:46

kaggle

About us   How it works   Find a competition   Host a competition   Blog   Help

is a platform for data prediction competitions that allows organizations to post their data and have it scrutinized by the world's best data scientists. See how it works.

Get started now to create or enter competitions:

Sign up

Hackathons as labour

TGS Salt Identification Challenge

Segment salt deposits beneath the Earth's surface
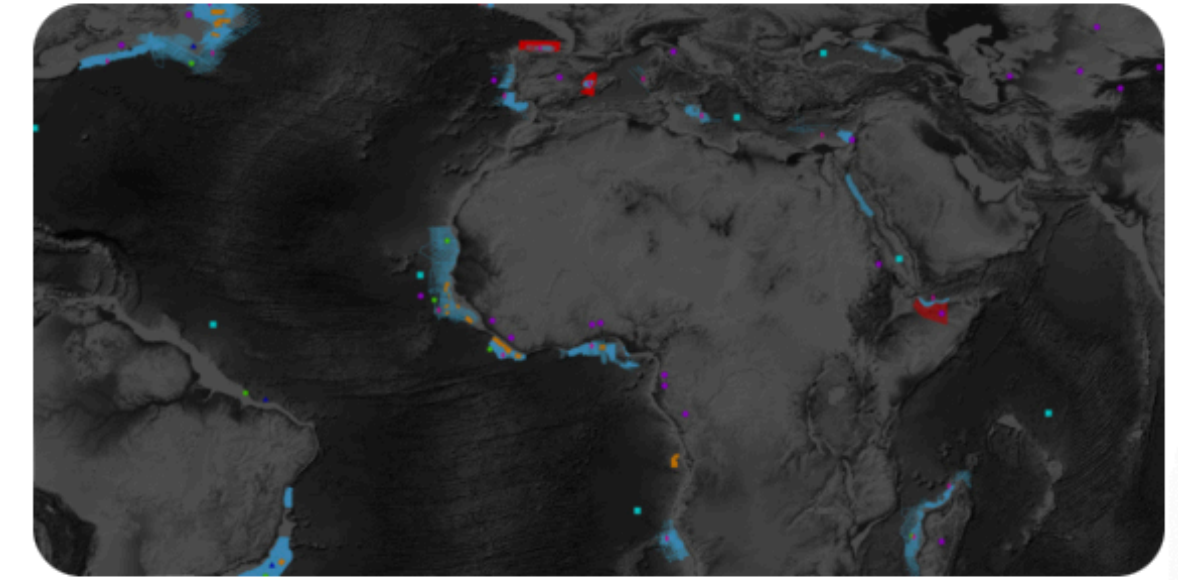
- "Several areas of Earth with large accumulations of oil and gas also have huge deposits of salt below the surface ... TGS is hoping Kaggle's machine learning community will be able to build an algorithm that automatically and accurately identifies if a subsurface target is salt ...

Hackathon as labour

- "... In addition to being accurate, however, trustworthy machine learning ... must also be *fair* ..."

- Is it fair to have trainees do hackathons?

  - Labour may not be adequately compensated

  - Output may not benefit trainee

Counterpoint

- Hackathons provide valuable training and networking opportunities

Synthesis

- Hackathons that benefit the commons/public good should be organized or promoted

Question

- To what extent are hackathons already benefiting the commons/public good?

Three years of Kaggle

- We examine three years of hackathons on kaggle, consisting of N = 125 hackathons with prize >$5k USD. Of these, 11 do not have traditional leaderboards, leaving a sample of N = 114 hackathons

- We categorize the hackathons into three groups:

  - Good. The hackathon benefits the commons/public good

  - Neutral. The hackathon may not benefit the commons/public good, but there are no obvious externalities

  - Other. The hackathon doesn't fall into one of the above two groups

Assigning values

- Good: The hackathon benefits the public good

Human Protein Atlas - Single Cell Classification

Find individual human cell differences in microscope images
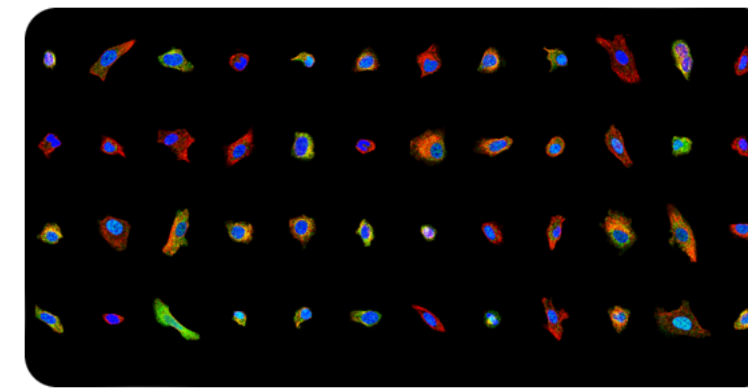
TensorFlow - Help Protect the Great Barrier Reef

Detect crown-of-thorns starfish in underwater image data

CDP - Unlocking Climate Solutions
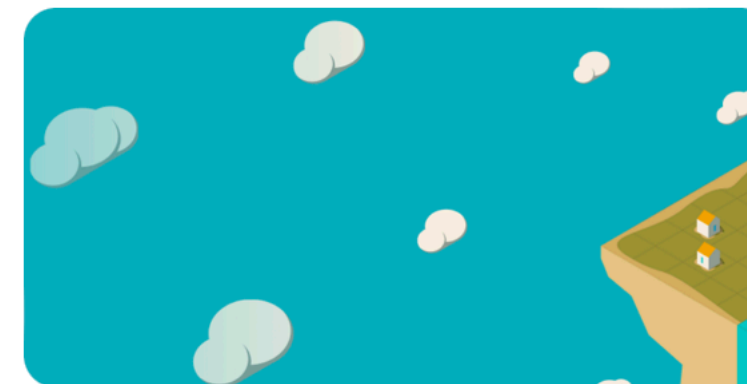
City-Business Collaboration for a Sustainable Future

Assigning values

- Neutral

Lux AI

Gather the most resources and survive the night!

AI Village Capture the Flag @ DEFCON

Hack AI! Collect flags by evading, poisoning, stealing, and fooling AI/ML

Assigning values

- Other

Indoor Location & Navigation
Identify the position of a smartphone in a shopping mall

G-Research Crypto Forecasting
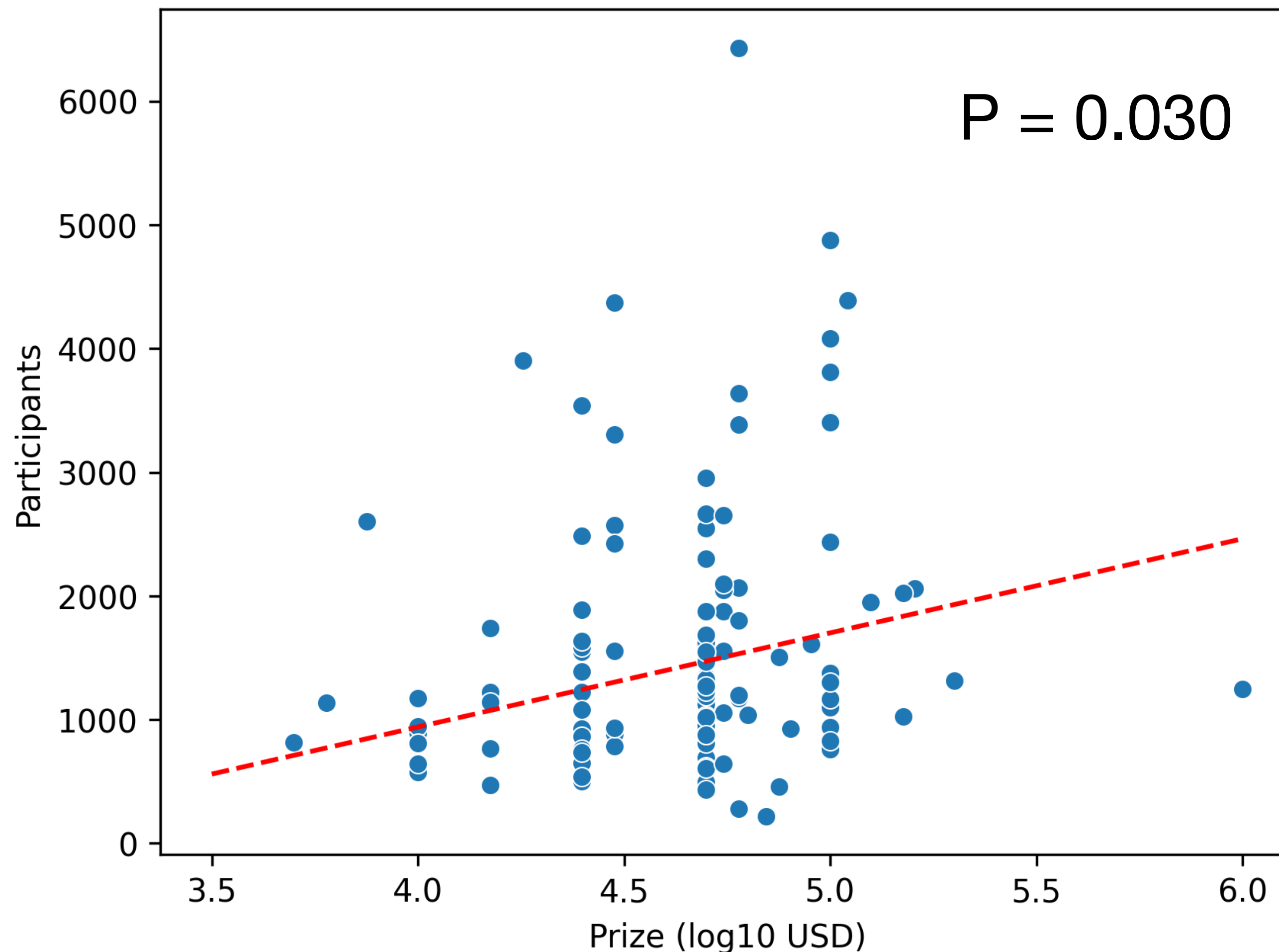Use your ML expertise to predict real crypto market data

U.S. Patent Phrase to Phrase Matching
Help Identify Similar Phrases in U.S. Patents

Motivation by prize

"Join the $1,000,000+ Vesuvius Challenge to resurrect an ancient library from the ashes of a volcano. In this competition you are tasked with detecting ink from 3D X-ray scans and reading the contents. Thousands of scrolls were part of a library located in a Roman villa in Herculaneum, a town next to Pompeii.

Summary statistics

Value	Number	Median prize	Mean participants	Mean participants (adjusted)
Good	56	50000	1475	-120
Neutral	26	50000	1770	181
Other	32	50000	1640	63

Statistical tests
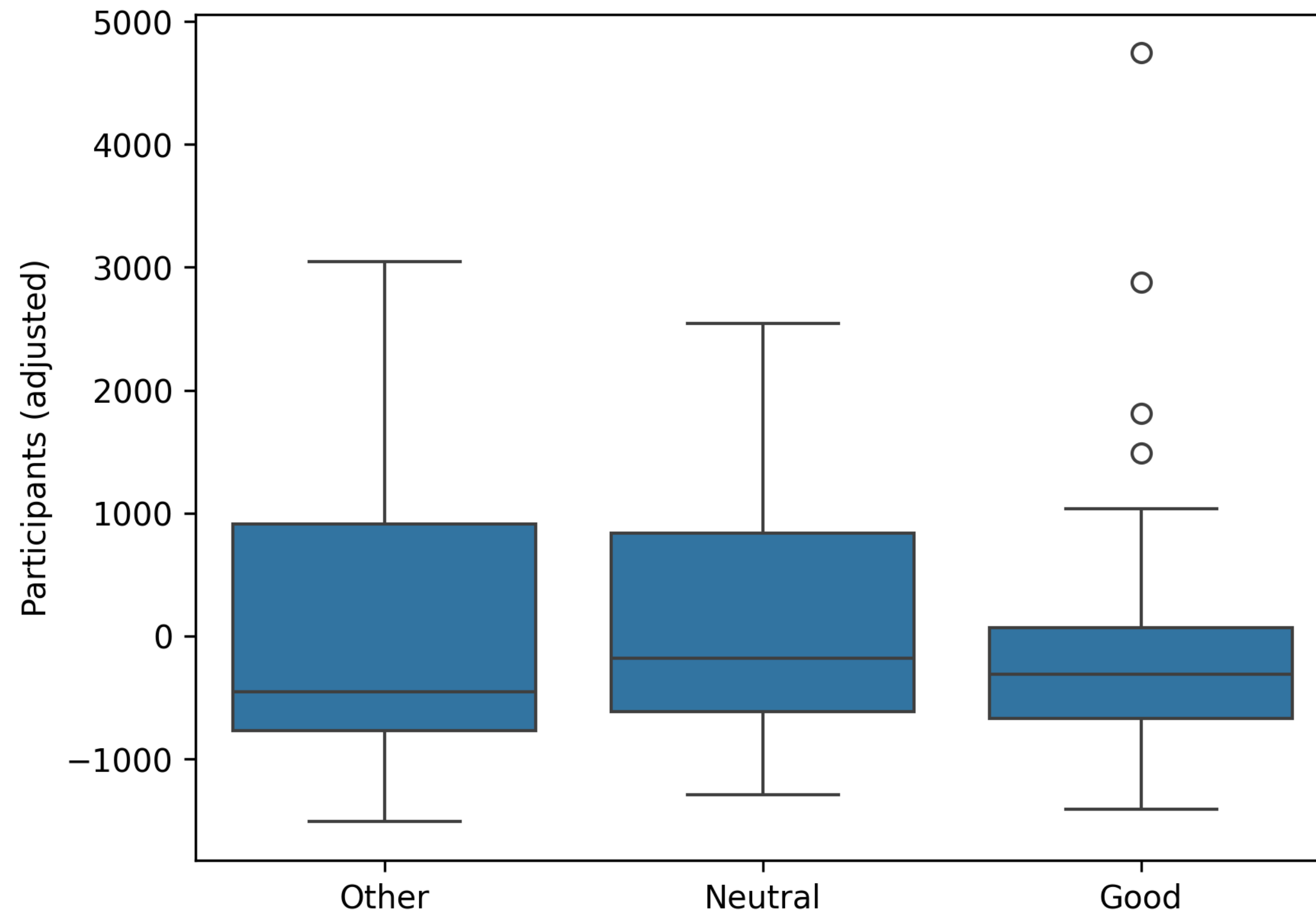
No significance. Pairwise p-values for two-sided t-test: 0.43, 0.22, 0.70

Binomial test for rate of good hackathons differing from 0.5 yields p-value 0.93

Errors

- No control for year, length of contest

- Unclear how well Participants is a proxy for labour

- Determination of value is sometimes arbitrary

Conclusion

- Results suggest:

  - No evidence of difference in participation between good and neutral or other

  - No evidence that good hackathons are happening at a lower (or greater) rate than neutral or other

  - Surprisingly, about half the hackathons on kaggle involve data for good, even ones with corporate organizers

TensorFlow - Help Protect the Great Barrier Reef
Detect crown-of-thorns starfish in underwater image data